Regresión lineal y series de tiempo Gerardo

Considere que b_0 y b_1 son las estimaciones de mínimos cuadrados para una nube de datos $\{(x_i, y_i)\}_{i=1}^n$. Ahora, definamos las siguientes cantidades

$S_{xy} := \sum_{i=1}^n (x_i - \bar{x_n})(y_i - \bar{y_n})$

$S_{xx} := \sum_{i=1}^n (x_i - \bar{x_n})^2$

$S_{yy} := \sum_{i=1}^n (y_i - \bar{y_n})^2$

$SCE := \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$

Demuestre que

$$SCE = S_{yy} - b_1 Sxy.$$

Demostración

$$SCE = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

como $b_0 = \bar{y_n} - b_1 \bar{x_n}$,

$$= \sum_{i=1}^{n} ((y_i - \bar{y_n}) + (b_1 \bar{x_n} - b_1 x_i))^2$$

$$= \sum_{i=1}^{n} ((y_i - \bar{y_n}) + b_1 (\bar{x_n} - x_i))^2,$$

desarrolando el cuadrado, tenemos

$$= \sum_{i=1}^{n} (y_i - \bar{y_n})^2 - 2b_1 \sum_{i=1}^{n} (y_i - \bar{y_n})(\bar{x_n} - x_i) + b_1 \sum_{i=1}^{n} (\bar{x_n} - x_i)^2,$$

recordemos que $b_1 = \frac{S_{xy}}{S_{xx}}$ y que $S_{xx} := \sum_{i=1}^{n}(x_i - \bar{x_n})^2$, por lo que

$$= \sum_{i=1}^{n}(y_i - \bar{y_n})^2 - 2b_1 \sum_{i=1}^{n}(y_i - \bar{y_n})(\bar{x_n} - x_i) + b_1 S_{xy},$$

por las definiciones de S_{yy} y S_{xy}, concluimos que

$$SCE = S_{yy} - b_1 Sxy.$$